

The Galaxy-Dark Matter Connection: A Cosmological Perspective

**Surhud More¹, Frank van den Bosch², Marcello Cacciato³,
Anupreeta More¹, Houjun Mo⁴, Xiaohu Yang⁵**

¹Kavli Institute for Cosmological Physics, University of Chicago, 933 East 56th Street, Chicago, IL 60637, USA

²Astronomy Department, Yale University , Box 208101, New Haven, CT 06520-8101, USA

³Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

⁴Department of Astronomy, University of Massachusetts, Amherst MA 01003-9305

⁵Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, the Partner Group of MPA, Nandan Road 80, Shanghai 200030, China

E-mail: surhud@kicp.uchicago.edu

Abstract. We present a method that uses observations of galaxies to simultaneously constrain cosmological parameters and the galaxy-dark matter connection (aka halo occupation statistics). The latter describes how galaxies are distributed over dark matter haloes, and is an imprint of the poorly understood physics of galaxy formation. A generic problem of using galaxies to constrain cosmology is that galaxies are a biased tracer of the mass distribution, and this bias is generally unknown. The great advantage of simultaneously constraining cosmology and halo occupation statistics is that this effectively allows cosmological constraints marginalized over the uncertainties regarding galaxy bias. Not only that, it also yields constraints on the galaxy-dark matter connection, this time properly marginalized over cosmology, which is of great value to inform theoretical models of galaxy formation. We use a combination of the analytical halo model and the conditional luminosity function to describe the galaxy-dark matter connection, which we use to model the abundance, clustering and galaxy-galaxy lensing properties of the galaxy population. We use a Fisher matrix analysis to gauge the complementarity of these different observables, and present some preliminary results from an analysis based on data from the Sloan Digital Sky Survey. Our results are complementary to and perfectly consistent with the results from the 7 year data release of the WMAP mission, strengthening the case for a true ‘concordance’ cosmology.

1. Introduction

Cosmology has reached an important cross-road in the last couple of decades, transitioning from a data-craved to a data-driven field of science. The concordance cosmological picture of a Universe dominated in energy density by dark energy and dark matter has emerged from a vast number of cosmological investigations (see Figure 2 in the contribution by P. J. E. Peebles in the current volume). Ordinary matter forms a fairly small component ($\sim 4.5\%$) of the energy density of the Universe, with most of it present in the form of intergalactic gas. The energy density of stars in galaxies has an extremely negligible contribution to the energy budget. However, unlike dark matter or dark energy, we can observe the star-light from galaxies directly, and use galaxies as tracers of the underlying matter density field to investigate the properties of the

Universe. Unfortunately, this connection between galaxies and (dark) matter is complicated by the fact that galaxies are biased tracers of the mass distribution. Although this ‘galaxy bias’ is generally considered a nuisance when trying to use galaxies to constrain cosmology, it also contains a wealth of information regarding galaxy formation. After all, it is the physics of galaxy formation that determines where, how and with what efficiency galaxies form within the dark matter density field. Therefore, ideally one would like to *simultaneously* solve for cosmology and galaxy bias. In this paper, we present a method that can do this, and show some preliminary results.

The overdensity of galaxies, δ_g , at a given position \vec{x} , is related to the overdensity of matter, δ_m , at that position, by a multiplicative term called the galaxy bias,

$$\delta_g(\vec{x}) = b_g \delta_m(\vec{x}), \quad (1)$$

and this implies that the power spectrum of the galaxy overdensity field on a particular scale is related to the matter overdensity power spectrum by

$$P_{gg}(k) = b_g^2(k) P_{mm}(k). \quad (2)$$

In general, the galaxy bias defined in the above manner is expected to be scale dependent [1]. However, on large scales, gravitation is the only relevant physics and galaxy bias is expected to be scale free and equal to a constant. The shape of the galaxy power spectrum on large scales, therefore, mimics the shape of the matter power spectrum. The investigations of cosmological parameters, in particular, the shape parameter $\Gamma = \Omega_m h$, have primarily focussed on large scale precisely for this reason [2, 3, 4]. However, it is also clear from the above equations that on large scales, the amplitude of the power spectrum, quantified by σ_8 , is perfectly degenerate with the galaxy bias, and large scale observations can only constrain the product $b\sigma_8$ very well [2]. The problem is further complicated by the well known result that brighter galaxies cluster more strongly than fainter galaxies, thus implying that the galaxy bias is also luminosity dependent [5].

It is crucial to understand why and how galaxies are biased with respect to the matter distribution in order to break the degeneracy between galaxy bias and the amplitude of the matter power spectrum. The matter distribution in the Universe collapses to form bound clumps of matter called halos. Since halos form preferentially at the peaks of matter density field, halos themselves are biased tracers of the underlying matter density field [6]. Galaxies form and reside within these halos of dark matter, and therefore they inherit the bias of their parent halos. Observations of the abundance of galaxies [7, 8], the clustering of galaxies on small scales [9, 10], the gravitational lensing signal due to the dark matter around galaxies [11, 12], and the kinematics of satellite galaxies around halos [13, 14, 15] can all provide important clues regarding this “galaxy-dark matter connection” (i.e., what galaxies resides in what halo). Using this information one can predict the galaxy bias, both as function of scale and as function of galaxy properties (e.g., luminosity). This allows one to break (some of) the degeneracies between galaxy bias and cosmology, and thus to use the observed distribution of galaxies to constrain cosmology.

Unfortunately, the clustering of galaxies is not sufficient to fully break degeneracies. This is easy to understand. A generic prediction of hierarchical formation scenarios is that more massive haloes are more strongly clustered. Hence, the observed clustering strength of a particular subset of galaxies (i.e., galaxies in a narrow luminosity bin) on large scales, is a direct measure for the characteristic mass of their dark matter haloes. However, different cosmologies predict different clustering properties of the dark matter haloes. Consequently, the galaxy-dark matter connection inferred from measurements of galaxy clustering are strongly cosmology dependent [26]. This degeneracy can be broken using additional, independent constraints on the galaxy-dark matter connection, such as those provided by satellite kinematics or galaxy-galaxy lensing.

In this paper, we demonstrate the strength and complementarity of a variety of galaxy observations to constrain cosmological parameters. In particular, we show how observations of galaxy abundances, galaxy clustering and galaxy-galaxy lensing, can be used to constrain cosmological parameters such as the matter density in units of the critical density, Ω_m , and the amplitude of the power spectrum of matter fluctuations, as characterized by σ_8 . We rely on the framework of the halo model to analytically predict these observations. The halo model assumes that all the dark matter in the Universe is partitioned over dark matter halos of different sizes and masses [16]. The abundance and clustering of these halos of dark matter is set by the underlying cosmological parameters, and this dependence has been well calibrated with the use of numerical simulations [17, 18]. A parametric form of how galaxies populate halos, called the halo occupation distribution function, can then be used to predict the abundance and clustering of galaxies using the abundance and clustering of halos [19]. In this paper, we use a Fisher matrix analysis to highlight the complementarity of using these different data sets, and we present some preliminary results from an analysis based on existing data.

2. Data

We use data from the main galaxy sample of the Sloan Digital Sky Survey (SDSS) [20, 21]. In particular, we use the galaxy luminosity function, $\Phi(L)$, of [22], the projected two-point correlation functions, $w_p(r_p)$, for six different luminosity bins and $0.2h^{-1}\text{Mpc} \lesssim r_p \lesssim 40h^{-1}\text{Mpc}$, obtained by [10], and the excess surface densities $\Delta\Sigma(r_p)$, for the same six luminosity bins but for $0.05h^{-1}\text{Mpc} \lesssim r_p \lesssim 2h^{-1}\text{Mpc}$ from [11]. The latter is proportional to the tangential shear induced by the mass distribution associated with the galaxies in question, and can be measured in the form of weak distortions of background galaxies due to weak gravitational lensing (galaxy-galaxy lensing). Our goal is to use a unified model that can describe all these observables in terms of a simple parametric model, and to use the existing data to simultaneously constrain cosmological parameters and halo occupation statistics.

3. Analytical framework

We use the conditional luminosity function (CLF) to specify the halo occupation distribution of galaxies [23]. The CLF, $\Phi(L|M)dL$, describes the average number of galaxies with luminosity $L \pm dL/2$ that reside in a halo of mass M , and consists of two components; one for central galaxies and the other for satellites. Motivated by results obtained from a large SDSS galaxy group catalogue [24], we assume that the CLF for central galaxies is described by a log-normal distribution with a logarithmic mean luminosity that depends on mass and a scatter which we assume to be mass-independent. The dependence of the logarithmic mean luminosity, \tilde{L}_c on halo mass is parameterized using four central CLF parameters, L_0 , M_1 , γ_1 and γ_2 , and is given by

$$\tilde{L}_c(M) = L_0 \frac{(M/M_1)^{\gamma_1}}{(1 + M/M_1)^{\gamma_1 - \gamma_2}}. \quad (3)$$

For the satellite component, we assume that it is well described by a Schechter-like function

$$\Phi_s(L|M) = \Phi_* \left(\frac{L}{L_s} \right)^{\alpha_s} \exp \left[- \left(\frac{L}{L_s} \right)^2 \right], \quad (4)$$

where the parameters, L_s , Φ_* and α_s are, in general, functions of halo mass. Guided by the results of [24] based on a galaxy group catalog, we assume that $L_s(M) = 0.562\tilde{L}_c(M)$ and that α_s is a constant, independent of halo mass. The function $\log \Phi_s$ is assumed to have a quadratic dependence on $\log M$, which is described by three free parameters, b_0 , b_1 and b_2 . Hence, the CLF, which describes the halo occupation distribution as function of galaxy luminosity, is described by a total of 9 free parameters.

In addition to specifying the luminosity dependence of the halo occupation distribution, we also need to specify the spatial distribution of galaxies in dark matter halos. Throughout we assume that central galaxies reside at the centers of their halos and that the satellite galaxies follow the density distribution of dark matter without any spatial bias. We have verified that this assumption has a negligible impact on our cosmological constraints.

Given the parameters of the central and satellite CLF, and the cosmological parameters which set the abundance and clustering of halos, we can predict all the observables that we wish to model. For example, the luminosity function simply follows from multiplying the average number of galaxies in a halo of mass M by the number density of halos of that mass, $n(M)$, and simply integrating this product over all halo masses [25],

$$\Phi(L) = \int \Phi(L|M) n(M) dM. \quad (5)$$

Similarly, the large scale bias of galaxies of luminosity L can be obtained by the following weighted average of the large scale bias of dark matter halos, $b_h(M)$, according to

$$b_g(L) = \frac{\int \Phi(L|M) b_h(M) n(M) dM}{\int \Phi(L|M) n(M) dM}, \quad (6)$$

Because of the page-limits of these proceedings, we cannot provide the detailed, analytical expresions that we use to calculate the observables $\Phi(L)$, $w_p(r_p)$, and $\Delta\Sigma(r_p)$ for a given model (i.e., cosmology plus CLF). These will be presented in van den Bosch et al. (2012, in preparation). We emphasize, though, that our implementation of the ‘halo model’ [16] properly accounts for (i) the scale dependence of halo bias, (ii) halo exclusion and (iii) residual redshift space distortions that can affect the determinations of galaxy bias [27]. Detailed tests using realistic mock galaxy catalogs indicate that our analytical model is accurate to better than 5 percent over the entire range of scales covered by the data.

Throughout we adopt a ‘standard’ flat Λ CDM cosmology (i.e., gravity is described by standard General Relativity, neutrino mass is neglected, initial power spectrum is a single power-law, and dark energy is modeled as Einstein’s cosmological constant), which is described by 5 cosmological parameters: the matter density parameter Ω_m , the baryon density parameter Ω_b , the hubble parameter h , the power law index n_s and the parameter σ_8 . Our goal is to constrain (subsets) of these cosmological parameters, fully marginalizing over the galaxy-dark matter connection as parameterized by our 9-parameter CLF model.

4. Results

4.1. Fisher forecasts

In this section we use the Fisher information matrix in order to gauge the accuracy with which constraints on the cosmological parameters Ω_m and σ_8 can be obtained given the current accuracy of the observables that we wish to model. Since we have three different observables, the luminosity function, galaxy-galaxy clustering and galaxy-galaxy lensing, we start by investigating how each of these different data sets contribute to our constraining power.

The different panels of Figure 1 show the 68, 95 and 99 percent confidence intervals that can be placed on the cosmological parameters, Ω_m and σ_8 under varying assumptions of prior information from the 7 year analysis of the WMAP mission [28]. The left-hand panel assumes uninformative priors on all of the cosmological parameters in our model. The dashed contours are used to indicate the confidence levels when we perform a joint analysis of the abundance of galaxies and the galaxy-galaxy lensing signal around them. The constraints are fairly weak, in particular, because the galaxy-galaxy lensing signal has only been measured on fairly small scales ($r_p \lesssim 2h^{-1}$ Mpc). This results in a number of degeneracies between the cosmological parameters

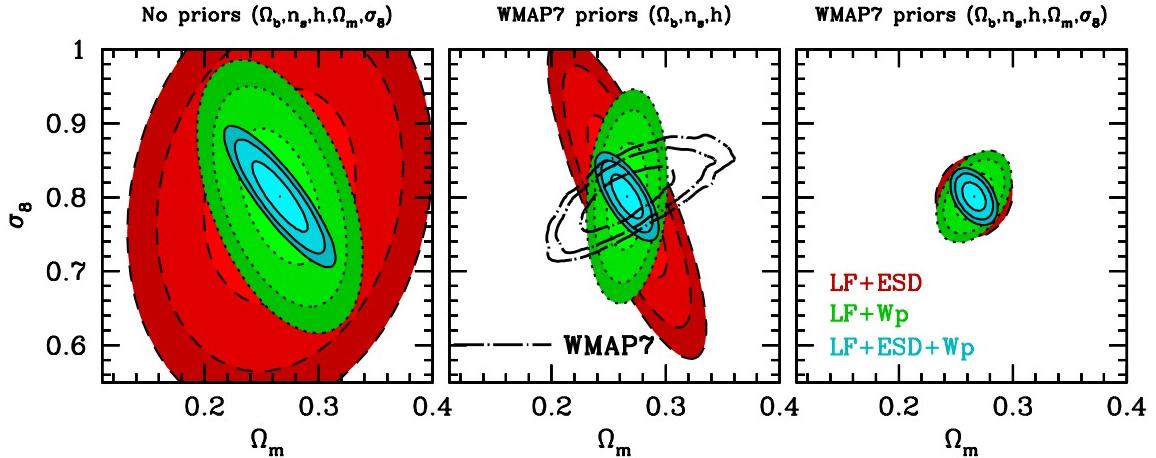


Figure 1. Fisher forecasts of 68, 95 and 99 percent confidence constraints on the cosmological parameters Ω_m and σ_8 when different combinations of the luminosity function (LF), the projected galaxy clustering (Wp) and the galaxy-galaxy lensing (ESD) data, as indicated in the legend, are analysed. Different panels show the effect of varying priors on cosmological parameters from analysis of the cosmic microwave background data. The dot-dashed contours in the middle panel show the constraints on Ω_m and σ_8 from the 7 year data release of the WMAP mission, and are shown for comparison. For a definitive version, see More et al. (2012), in preparation.

and the CLF parameters such that Ω_m and σ_8 are only weakly constrained. The dotted contours show the confidence contours obtained by combining the luminosity function with the galaxy-galaxy clustering data. The constraints are significantly tighter, and the improvement is largely due to the addition of information on intermediate scales ($2h^{-1}\text{Mpc} \lesssim r_p \lesssim 40h^{-1}\text{Mpc}$). Finally the solid contours show the result of a joint analysis of all three observables. Even in the absence of prior information, this joint analysis breaks a number of degeneracies that are present in our model. The resulting cosmological constraints are competitive with the existing constraints on these parameters, demonstrating the potential power of this method.

The middle panel of Figure 1 shows the effect of adding prior information on the secondary cosmological parameters, Ω_b , n_s and h from the 7 year WMAP results (hereafter WMAP7). Notice how the addition of prior information can flip the directions of degeneracies (compare the dashed contours in the left-hand and the middle panels). The degeneracy between Ω_m and σ_8 from our analysis is such that it runs perpendicular to the WMAP7 constraints (shown by the dot-dashed contours). Adding these WMAP7 constraints as additional priors on Ω_m and σ_8 , further improves the constraints, as shown in the right-hand panel.

The cosmological constraints presented above are also competitive with those obtained from studies of the abundance of galaxy clusters as a function of redshift. These galaxy clusters are detected either via their X-ray emission [29], or as overdensities in optical galaxy catalog [30], or with the Sunyaev-Zel'dovich effect [31, 32] and require extensive followup to calibrate the cluster mass-observable relationship. Measurements of cosmic shear have been used to obtain cosmological constraints, albeit weaker, on Ω_m and σ_8 [33, 34, 35, 36]. Cosmological constraints have also been recently obtained by analysing galaxy clustering and the mass-to-number ratio on cluster scales by [37]. It is important to note that all of these methods have very different systematics and are complementary to each other and our method: they are all part of a network of tests designed to validate the Λ CDM paradigm.

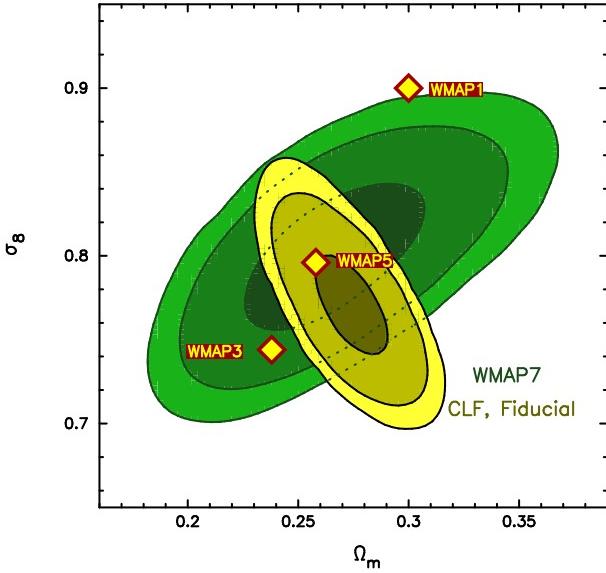


Figure 2. 68, 95 and 99 percent confidence limits on the cosmological parameters Ω_m and σ_8 from our analysis (shown in chrome yellow) compared with the confidence limits obtained by the analysis of the seven year data from the cosmic microwave background experiment WMAP (shown in green). For a more definitive version, see Cacciato et al. (2012, in preparation).

4.2. Cosmological constraints

We have carried out a joint analysis of all three observables, the luminosity function, the projected galaxy clustering and the galaxy-galaxy lensing signal, and obtained the posterior distribution of our model parameters given these data. We use a Monte-Carlo Markov chain to sample from posterior probability distribution of the parameters. For our fiducial analysis, we impose priors on the secondary cosmological parameters Ω_b , n_s and h and completely uninformative priors on the parameters Ω_m and σ_8 . Our model is able to fit the data sufficiently well with χ^2 per degree of freedom of the order of 2. The fits to the data will be presented in Cacciato et al. (2012, in preparation). Preliminary results of our analysis, in the form of 68, 95 and 99 percent confidence contours, are shown in Fig.2 and compared to the WMAP7 results.

There are two points worth making. First of all, the constraints obtained from our analysis are in remarkably good agreement with the WMAP7 results, even though we have used no prior information on Ω_m and σ_8 . The WMAP7 results are based on observations of the microwave background at a very early time in the Universe ($z \sim 1080$) and primarily rest on the physics of perturbations that can be treated with the help of linear perturbation theory. The results from our analysis derive from galaxy observations at redshift $z \sim 0.1$ and are obtained by modelling extremely non-linear scales, properly marginalizing over the uncertainties related to galaxy bias (i.e., the galaxy-dark matter connection). The agreement in cosmological constraints obtained from these two completely disjunct analyses is extremely striking and provides strong support for the notion of a true ‘concordance’ cosmology: clearly Λ CDM provides an excellent description of data over a large range of scales and cosmic epochs. Secondly, the constraints obtained from our analysis are both competitive with and complementary to those obtained by the WMAP analysis. This is also in agreement with the complementarity expected from the Fisher forecast presented in the previous subsection.

5. Summary

To summarize, observations of galaxies are an excellent way of probing the underlying matter distribution in the Universe and thereby obtaining precise constraints on the cosmological model. We have shown that a joint analysis of the abundance of galaxies (characterized by the galaxy luminosity function), the clustering of galaxies (characterized by the projected two-point correlation functions), and the clustering of dark matter around galaxies (characterized by galaxy-galaxy lensing) can be a useful way to constrain the cosmological parameters. We have

modelled each of these observations in the analytical framework of the halo model. The halo occupation distribution of galaxies in our model was specified by the parametric CLF model.

Using a Fisher matrix analysis, we have shown that the cosmological information contained in the three observables described above is complementary to each other and a joint analysis of these datasets is able to break a number of degeneracies between the CLF parameters and the cosmological parameters. We followed up our Fisher forecast results, by constraining our model parameters using the actual data. We have shown that the resulting constraints on the cosmological parameters Ω_m and σ_8 are in remarkable agreement with constraints from the analysis of the WMAP data. This is yet another jewel in the crown of the Λ CDM model, which continues to reign king.

We are currently exploring the use of our method to constrain extensions of the Λ CDM model that include cosmological parameters such as the neutrino density parameter, the dark energy equation of state, non-gaussianity in the initial density fluctuations and modifications to gravity. However, a large amount of work is still required in order to calibrate the predictions of the abundance and clustering of dark matter haloes in these alternative cosmologies. Although many such calibrations for the extended Λ CDM already exist in the literature, the halo mass definitions used in these calibrations are often not suitable for use in the halo model [38]. We expect to address these issues and extensions of our methods in future work.

Finally, an interesting by-product of our analysis is a detailed, statistical description of the galaxy-dark matter connection, as parameterized by the CLF, fully marginalized over cosmological uncertainties. This galaxy-dark matter connection is the outcome of a large number of poorly understood astrophysical processes, such as the formation of stars and the regulation of galaxy growth by feedback, that shape the formation and evolution of galaxies in our Universe. By constraining the CLF we are therefore constraining the integral effect of these (and other) processes. Hence, the constraints on the CLF parameters obtained from our analysis will be of great value to inform theoretical models of galaxy formation (both semi-analytic models and direct numerical simulations).

Acknowledgements

This research has spanned more than three years during which the affiliations of the different authors have changed. SM, FvdB and MC would like to acknowledge support from the Max Planck Institute for Astronomy and the University of Utah during the partial conduct of this research. We also thank Jeremy Tinker, Alexie Leauthaud, Martin White, Andrey Kravtsov, Eduardo Rozo, Matt Becker, Yin Li and Wayne Hu for many interesting discussions and possible extensions of this research work.

References

- [1] Cacciato M., Lahav O., van den Bosch F. C., Hoekstra H., Dekel A., 2012, preprint (arXiv:1203.2616)
- [2] Tegmark M., et al., 2004, ApJ, 606, 702
- [3] Percival W. J., et al., 2007, ApJ, 657, 645
- [4] Reid B. A., et al., 2010, MNRAS, 404, 60
- [5] Norberg P., et al., 2001, MNRAS, 328, 64
- [6] Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
- [7] Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
- [8] Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
- [9] Zehavi I., et al., 2005, ApJ, 630, 1
- [10] Zehavi I., et al., 2011, ApJ, 736, 59
- [11] Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715
- [12] Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, ApJ, 744, 159
- [13] More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, MNRAS, 392, 801
- [14] More S., van den Bosch F. C., Cacciato M., 2009, MNRAS, 392, 917
- [15] More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210

- [16] Cooray A., Sheth R., 2002, PhR, 372, 1
- [17] Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
- [18] Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
- [19] Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, MNRAS, 394, 929
- [20] York D. G., et al., 2000, AJ, 120, 1579
- [21] Abazajian K. N., et al., 2009, ApJS, 182, 543
- [22] Blanton M. R., et al., 2003, ApJ,
- [23] Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057
- [24] Yang X., Mo H. J., van den Bosch F. C., 2008, ApJ, 676, 248
- [25] van den Bosch F. C., et al., 2007, MNRAS, 376, 841
- [26] van den Bosch F. C., Mo H. J., Yang X., 2003, MNRAS, 345, 923
- [27] More S., 2011, ApJ, 741, 19
- [28] Komatsu E., et al., 2011, ApJS, 192, 18
- [29] Vikhlinin A., et al., 2009, ApJ, 692, 1060
- [30] Rozo, E., Wechsler, R. H., Rykoff, E. S., et al. 2010, ApJ, 708, 645
- [31] Sehgal N., et al., 2011, ApJ, 732, 44
- [32] Benson, B. A., de Haan, T., Dudley, J. P., et al. 2011, arXiv:1112.5435
- [33] Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116
- [34] Massey, R., Rhodes, J., Leauthaud, A., et al. 2007, ApJS, 172, 239
- [35] Lin, H., Dodelson, S., Seo, H.-J., et al. 2011, arXiv:1111.6622
- [36] Huff E. M., Eifler T., Hirata C. M., Mandelbaum R., Schlegel D., Seljak U., 2011, preprint (arXiv:1112.3143)
- [37] Tinker J. L., et al., 2012, ApJ, 745, 16
- [38] More S., Kravtsov A. V., Dalal N., Gottlöber S., 2011, ApJS, 195, 4